UNITED STATES DEPARTMENT OF AGRICULTURE

FOREST SERVICE

A BRIEF RESUMÉ OF THE YATES' LECTURES.

By
B. B. Day.

M—5123

U. S. GOVERNMENT PRINTING OFFICE:

8—7417

560084

UNITED STATES DEPARTMENT OF AGRICULTURE

FOREST SERVICE

ADDRESS REPLY TO
THE FORESTER
AND REFER TO

RS
ME

December 13, 1937.

Director,
California Forest & Range Exp. Station,
Berkeley, California.

Dear Sir:

A number of requests have come from the field for mimeographed copies of the Yates lectures and conferences on statistical methods given before the Department Graduate School October 28 to 30. In the absence of any such record having been made it was believed some of the men might be interested in learning what was the general nature of the subject matter covered. For this reason, the attached brief resume was prepared by Miss Day from her personal notes. As such this necessarily must be extremely sketchy as no attempt was made to supplement the notes taken.

Very truly yours,

I. T. Haig

I. T. HAIG,
Acting Chief, Division of Silvics.

Enc. (10)

A BRIEF RESUMĒ OF THE YATES' LECTURES* OCTOBER 28 TO 30, 1937

Those men who have kept up with writings on modern statistical theory would have found very little that was new in these conferences. However, they were most stimulating and worthwhile. One was impressed with the fact that much of the subject matter Mr. Yates presented had been developed by him from actual field experience, the necessity of finding suitable statistical tools for the solution of practical problems of the utmost importance, and that these same tools had been tested by practical application and found to be good. His approach at all times was simple and informal, introducing only the minimum of mathematical theory and language. This was for the benefit of non-mathematical persons to whom the particular subject matter was more or less unfamiliar, but who were interested in its application to their special problems. In the case of the development of new designs he stated that always the pressure has been from the practical agronomist and not the theorist.

*Lectures and conferences held at the graduate school, U. S. Dept. of Agriculture by Frank Yates, Chief Statistician, Rothamsted Experiment Station, Harpenden, England.

Lecture I. Principles underlying the design of factorial (complex) experiments.

The subject matter of this lecture with the exception of some very recent developments presented may be found in Fisher's Design of Experiments.

Factorial Design, originally called Complex, essentially involves the inclusion of more than one factor, very simple and in fact quite old. The ordinary simple factorial designs use 2 or more factors having all levels.

Figure 1, an experiment on potatoes, was introduced here, 32 plots with 4 blocks of eight plots, all treatments appearing in each block, hence, 4 replications of each treatment. It included all combinations of 3 factors with two levels of each.

Nitrogen (n)
Potash (k)
Dung (d)
} with two levels of each

No treatment called (1)
Total number of treatments = 8 from 2 x 2 x 2 = 8
The treatment combinations indicated on diagram
The last line shows mean yields
Consider the interpretation of these mean yields:

Response to dung: n and k absent, d - (1) = 8.6 - 2.8 = 5.8
n present k absent, nd - n = 9.4 - 2.8 = 6.6
k " n " , kd - k = 11.2 - 7.5 = 3.7
n and k present, nkd - nk = 12.1 - 8.1 = 4.0
4/ 20.1
5.025

Mean response of dung, D = 5.025
Likewise of potash, K = 3.8
and Nitrogen, N = 0.6
} Main effects

Precision is high since in each case it is the mean of 16 plots minus 16 plots.

With these 32 plots the same precision is gained as if 16 replications of nitrogen alone for two levels had been used. Hence, one advantage of complex experiments is large gains in precision. Also we get even more important information the effects in response when other factors are present. For example:

Dung with absence of potash $\frac{5.8 + 6.6}{2} = 6.2$

" " presence " " $\frac{3.7 + 4.0}{2} = 3.8$

Difference in interaction of dung and potash -2.4

This also the mean of 16 plots minus the mean of 16 other plots (This is only true for 2 x 2 x 2).

For convenience the conventional factor of 1/2 is introduced, hence - 2.4 = 2 D.K and DK = -1.2

(Over)

Then also:

Main effect with K absent 5.0 - (-1.2) = 6.2
K present 5.0 + (-1.2) = 3.8

From the above simple one, more difficult ones may be introduced.

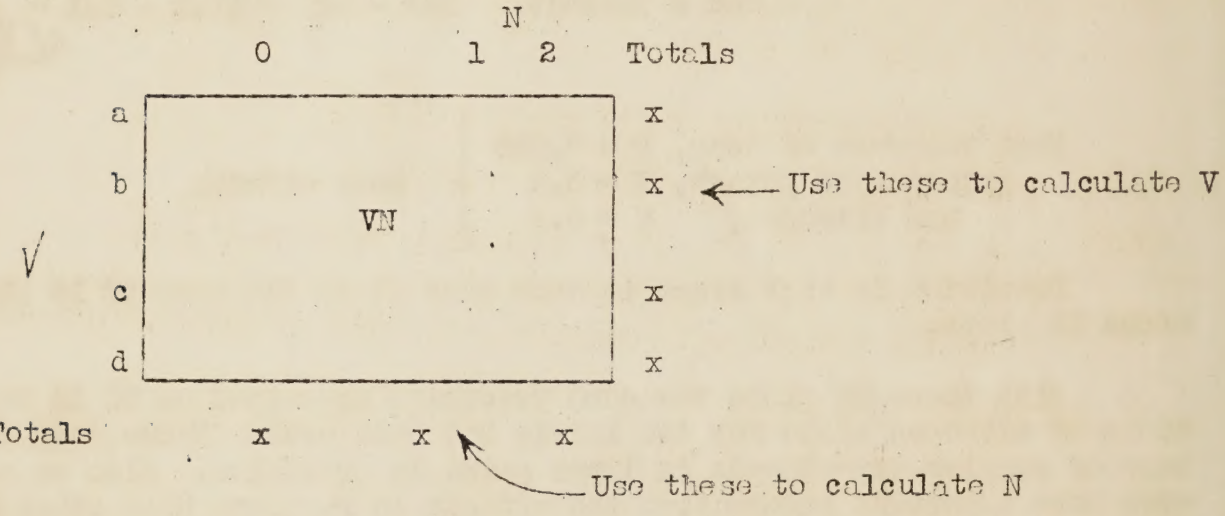
Consider the simple physical problem given a weighing machine of very fine order for which a zero correction is wished, and seven objects separately (See Figure 2). Instead of making eight weighings if weight a is desired it may be expressed as follows:

$$(W_1 + W_2 + W_3 + W_4) - (W_5 + W_6 + W_7 + W_8) = 4a$$

Hence, this dodge increases precision 4 times. In this case you know there are no interactions.

In an experiment of factors N and V, with the first having 3 levels and the second 4, there will be 12 treatments.

Then the set-up is as follows:



The inclusion of additional factors increase the number of plots rapidly.

Figure 3 is a 3 x 3 x 3 design showing sets of number, the 27 treatments being divided into 3 parts. If considering only two factors, each block is a replication. With a random arrangement on the ground the degrees of freedom are as follows:

N - 2	NP - 4	
P - 2	NK - 4	NPK - 8
K - 2	PK - 4	

The treatment NPK will be confounded with blocks. Actually confounded only partially for 2 degrees of freedom. If different combinations are used with 4 replications 3/4 of the information for NPK will be given. Always possible to do this with 2's and 3's but not with other numbers of levels.

Some other developments in designs are noted below.

Figure 4 is a coded system for experiment of $\underbrace{3 \times 3}_{9} \times \underbrace{3 \times 3}_{9}$, where the first number represents a latin square of nine. Likewise the second figure another latin square. In this design 81 blocks with information on 3 levels of 4 factors. Columns may be eliminated and get full precision of a 9 x 9 latin square. These are called quasi-latin squares.

Another development is an extension of confounding to non-factorial designs. With 81 varieties there are 9 groups of 9 varieties but not possible to compare varieties in one block with those in others. To overcome this cut across grouping. The precision is $\frac{p+1}{p+3} \frac{10}{12} \frac{5}{6}$ called quasi-factorial or lattice. Arrange by taking rows as one set of blocks and columns as other set.

If the simple problem is given to compare in pairs, as in twins, a, b, c, d, e, the standard method would be

a with b
a " c
a " d
a " e

S E of the difference of a - b = $\sqrt{\frac{1}{3}}\,\sigma$

" " " " " b - c = (b-a) - (c-a)

and " " " " " whole lot would be $\sqrt{\frac{2}{3}}\,\sigma$

Suppose instead all possible pairs were included. Others are

b - c
b - d e - d
b - e e - e d - e

Two replications of the first gives 5 of the second; same precision for each pair.

Semi-latin squares. Time did not allow for more than a brief mention of this type of design. Mr. Yates called attention, however, to a new publication, intended as a supplement to Fisher's "Design of Experiment", which discusses in some detail all the foregoing designs and others for which there was not sufficient time in this lecture. This publication is Emperial Bureau of Soil Science Technical Communication No. 35, price 5 shillings. The Design and Analysis of Factorial Experiments by F. Yates.

Fig. 1

A 2 x 2 x 2 EXPERIMENT ON POTATOES

Plan and Yields in Lbs.

Block I 2296				Block II 2291			
nk 291	kd 398	d 312	nd 373	kd 407	d 324	k 272	nk 306
(1) 101	k 265	n 106	nkd 450	n 89	nkd 449	nd 338	(1) 106
d 323	(1) 87	nd 324	kd 423	nd 361	nk 272	n 103	d 324
nk 334	k 279	n 128	nkd 471	k 302	(1) 131	nkd 437	kd 445
Block III 2369				Block IV 2375			

Yields of the Different Combinations of Treatments (Tons Per Acre).

(1)	n	k	nk	d	nd	kd	nkd	Mean
2.8	2.8	7.5	8.1	8.6	9.4	11.2	12.1	7.8

Dung versus no dung

		Means	Difference
n and k absent 8.6 - 2.8 =	5.8)	6.2)	
n present, k absent =	6.6)	)	
		)	-2.4 = 2 D x K
n absent, k present =	3.7)		
n and k present =	4.0)	3.8)	
	4/20.1		

Mean response of dung D = 5.025) Main Effects
K = 3.8)
N = 0.6)

D x K = -1.2) Interaction between two factors
N x K = +0.2)
N x D = +0.3)

N x D x K = -0.1

S.E.S. Single plot. ±0.50 = 6.2%
Main effects and interactions = ±0.18

n = sulphate of ammonia
k = sulphate of potash
d = dung

Fig. 2. A SCHEME FOR WEIGHING LIGHT OBJECTS ON A DIAL SCALE

(Zero connection to be determined)

Weighing	a	b	c	d	e	f	g	Wt.
1	+	+	+	+	+	+	+	W_1
2	+	+	+					W_2
3	+			+	+			W_3
4	+					+	+	W_4
5		+		+		+		W_5
6		+			+		+	W_6
7			+	+			+	W_7
8			+		+	+		W_8

$(W_1+W_2+W_3+W_4) - (W_5+W_6+W_7+W_8) = 4a$
$(W_1+W_2+W_5+W_6) - (W_3+W_4+W_7+W_8) = 4b$
$(W_1+W_2+W_7+W_8) - (W_3+W_4+W_5+W_6) = 4c$
$(W_1+W_3+W_5+W_7) - (W_2+W_4+W_6+W_8) = 4d$
$(W_1+W_3+W_6+W_8) - (W_2+W_4+W_5+W_7) = 4e$
$(W_1+W_4+W_5+W_8) - (W_2+W_3+W_6+W_7) = 4f$
$(W_1+W_4+W_6+W_7) - (W_2+W_3+W_5+W_8) = 4g$

Fig. 3 CONFOUNDING IN A 3 x 3 x 3 EXPERIMENT

Block	I			II			III		
Factor	N	P	K	N	P	K	N	P	K
	0	0	0	0	0	2	0	0	1
	1	0	1	1	0	0	1	0	2
	2	0	2	2	0	1	2	0	0
	0	1	2	0	1	1	0	1	0
Level	1	1	0	1	1	2	1	1	1
	2	1	1	2	1	0	2	1	2
	0	2	1	0	2	0	0	2	2
	1	2	2	1	2	1	1	2	0
	2	2	0	2	2	2	2	2	1

Fig. 4.

3 x 3 x 3 x 3 QUASI-LATIN SQUARE

11	29	35	48	54	63	76	82	97
28	34	13	56	62	47	81	99	75
36	12	27	61	49	55	98	74	83
45	51	69	73	88	94	17	26	32
53	68	44	87	96	72	25	31	19
67	46	52	95	71	89	33	18	24
79	85	91	14	23	38	42	57	66
84	93	78	22	37	16	59	65	41
92	77	86	39	15	21	64	43	58

Each number indicates a treatment combination of the four factors. The first digit indicating the combination of the first two factors and the second digit the combination of the third and fourth factors according to the following scheme:

	1	2	3	4	5	6	7	8	9
Level of the first factor	0	1	2	0	1	2	0	1	2
Level of the second factor	0	0	0	1	1	1	2	2	2

All main effects and interactions between two factors are clear of row and column effects, hence, the precision attained on these comparisons is that of a 9 x 9 Latin Square.

Lecture II. Contrasts between the methods of correlation and regression.

What is meant when it is said variables are correlated?

What is the underlying principles involving regression and correlation coefficients and what is the difference between them?

A diagram with two variables as

x	y
x_1	y_1
x_2	y_2
⋮	⋮

was introduced which showed the distribution when high values of x go with high values of y, this diagram a set of elliptical curves.

Best known measure for such a distribution is the correlation coefficient.

The variance of x is σ_x^2 or V(x)

" " " y " σ_y^2 " V(y)

Also the covariance of x and y expressed as Cov (xy)

and correlation coefficient $r = \frac{Cov(xy)}{V(x)V(y)} = \frac{S(x-\bar{x})(y-\bar{y})}{\sqrt{S(x-\bar{x})^2 S(y-\bar{y})^2}}$

This is a ratio and hence unaffected by scale; a special virtue where scale has no physical meaning.

Here a figure was introduced to illustrate the meaning of correlation.

Line drawn through mean of every set of values of y for each constant x. This line called the regression of Y on x. $Y = a + bx$ may be used in predicting y given x.

If b is the regression coefficient then $Y = \bar{y} + b(x-\bar{x})$

estimated $b = \frac{Cov\ (xy)}{V(x)}$ or $\frac{y}{x} r$.

If variances are equal b = r.

The regression line passes through points of vertical tangents.

Another aspect is the reduction of variance in y by use of regression coefficient. This reduction in variance in $y = V(y)\ \sqrt{1-r^2}$

A table of the significance of correlation, r, depends on the number of observations.

The significance of b may be determined given the standard error of b. This is easier to comprehend.

Usually the variation not the same for y and x as in

$y = a + bx + cx^2 + \ldots$, or $Y = f(x)$. Here correlation coefficient breaks down - meaning nothing, while regression continues to have meaning.

The following special case was introduced: A problem involving wheat yield and acreage and it was of course known that zero yield for zero acres. Do not attempt to bring regression line to this point. Use instead the expression "within range of observations."

Another objection to use of correlation coefficient is that errors in y will inflate the variance of y, while such errors do not effect b. There is no physical concept for the correlation coefficient and corrections must be made for errors in y. Correlation coefficient is upset by errors in either x or y or both.

Selection of independent variables will cause variations in correlation coefficient. Hence, there is never any justification for computing correlation coefficient where x is selected or not entirely random. An example was given where this was done with vitiated results.

Partial correlation and regression are what one gets with three or more variables. The joint distribution is specified by

$V(x)$, $V(y)$, $V(z)$, $Cov\,(xy)$, $Cov(xz)$, $Cov(yz)$.

and $Z = c + b_1 x + b_2 y$ or $Z = f(xy)$

There are also partial correlation and regression coefficients as

$$r_{zx.y},\ r_{zy.x} \quad \text{and} \quad b_1 = r_{zx.y} \frac{\sigma_z}{\sigma_x} \sqrt{\frac{1 - r^2_{yz}}{1 - r^2_{xy}}}, \text{ etc.}$$

It is often believed that it doesn't matter which is used partial correlation coefficient or partial regression coefficient. Relations between the two is not at all simple. Same relation as existed when variances were equal for two variables not true for partial correlation. One should use regression coefficients rather than correlation coefficients and compute their S.E.'s.

A number of examples from scientific literature were quoted wherein wrong use had been made of partial correlation coefficients.

Two other important concepts.

1. Intra-class correlation - better to go over to analysis of variance.

2. Multiple correlation.

$z = b_1 x + b_2 y$ correlate observed and predicted,

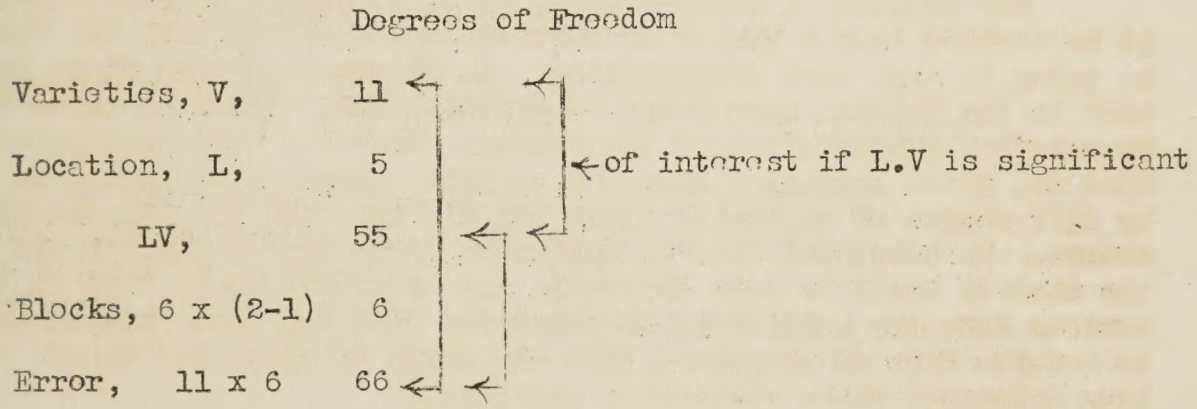
$(1 - r^2)$ portion removed. Here again Mr. Yates strongly advocated analysis of variance to be the better method.

To determine which is more important x or y in relation to z in $z = b_1x + b_2y$. Use $z = b_1'x$ and $z = b_2'y$ and compare total amount of variance removed in each case. If x and y are the same kind of data make a direct comparison between b_1 and b_2.

An account of the group conferences must of necessity be very general since these were mostly extempore discussions of a number of complex experiments now being conducted. Mr. Yates was able to clear up many obscure points in the minds of some regarding designs and the technique of analyses and to offer timely suggestions and warnings of possible pitfalls which might vitiate conclusions from such experiments.

A long time rotation wheat experiment on the use of various fertilizers in process at Rothamsted was described by Mr. Yates. Already valuable information is being gained although the experiment has a number of years yet to go.

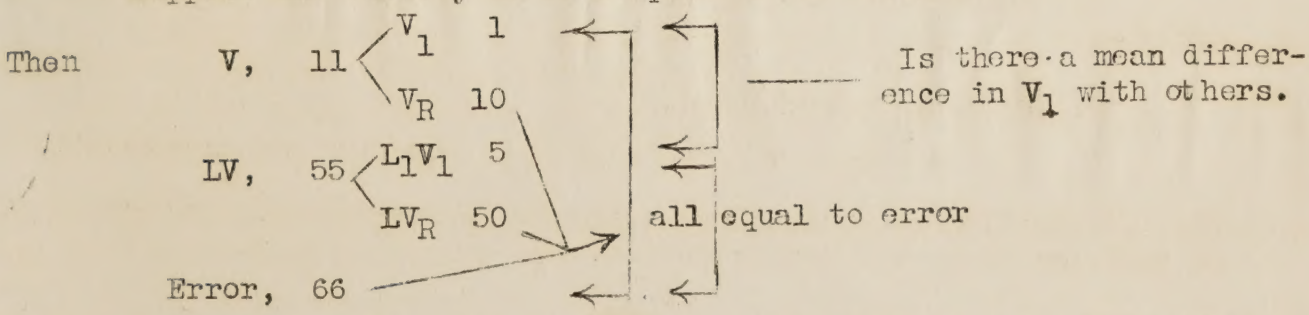
A regional cotton wilt variety fertilizer study in 12 locations for a 3-year period involving 12 varieties with three levels of potash and 3 replications at each location was discussed at length and the degrees of freedom by factors outlined. Mr. Yates called attention to the fact that while the pooling of the error term (in this case 70 x 35) from different locations resulted in great strength it might not be justifiable. He illustrated his meaning by an example. Suppose 12 varieties at 6 places with 2 replications at each.



Suppose locations are random of all locations are varieties significant answered by comparison of V and Lv?

Question of testing consistency of varietal differences.

Suppose one variety to be compared with others.



L_1V_1 may be very different from LV_R then pooling may not be desirable.

Suppose 1st variety is up then analyze be setting $V_1 - V_r$ for 6 places and test by t. If V_2 is a little different one could take $V_2 - \overline{V}_{r-1}$, etc.

By comparing each with mean of all others is a better basis than each with the mean of all.

$$(V_1 - \overline{V}_{2-12}) = (V_1 - \overline{V}_{1-12}) \frac{12}{11}$$

Precision of comparison of one variety with standard where it has more replication is increased over the precision of comparison of two varieties. Hence, if there should exist a standard it is well to increase number of replication of it.

Suppose errors are different then test these for significance. It is well to start with separate errors and later pool if not significantly different.

Mr. Yates described the work being done by the British Commission of Forestry in making a survey of forest resources for which he is now acting as Consultant. For such a survey he considered the best solution would be to make a random selection of type sections and then take intensive grid sample of each of these. For British conditions the circular $\frac{1}{10}$ A plot seemed adequate with about a 1/2% of total area.

For cruising a smaller area, e.g., a county, he recommended that it be divided into a number of sections and not less than two random plots be taken in each such subdividions. An alternate method might be the one used in the British Survey and described above, a grid of plots in randomly selected subdivisions of the county. When the grid arrangement was used Mr. Yates intimated that the sampling error might better be computed by differences of adjoining plots but did not give the details of this procedure. He advocated that in countries where forest regions are so variable that it would be most desirable for research to be done in these various forestry conditions to determine what the best methods are, such to involve type of sampling, size and shape of plots and degree of sampling necessary for a particular precision.

B. B. DAY